ARI Research Note 92-29

A User's Introduction to Determining Cost-Effective Tradeoffs Among Tank Gunnery Training Methods

John E. Morrison and R. Gene Hoffman
Human Resources Research Organization

for

Contracting Officer's Representative
Donald F. Haggard

Field Unit at Fort Knox, Kentucky
Barbara A. Black, Chief

Training Systems Research Division
Jack H. Hiller, Director

April 1992

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

MICHAEL D. SHALER
COL, AR
Commanding

Research accomplished under contract for the Department of the Army

Human Resources Research Organization

Technical review by

Scott E. Graham

Accession For		
NTIS GRA&I		☑
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1		

QUALITY INSPECTED 4

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1992, April	Final Report Feb 90 - Sep 91

4. TITLE AND SUBTITLE
A User's Introduction to Determining Cost-Effective Tradeoffs Among Tank Gunnery Training Methods

5. FUNDING NUMBERS
DAHC-35-89-C-0009
62785A
790
3205
C1

6. AUTHOR(S)
Morrison, John E.; and Hoffman, R. Gene

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Human Resources Research Organization
1100 S. Washington Street
Alexandria, VA 22314

8. PERFORMING ORGANIZATION REPORT NUMBER
--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
Training Systems Research Division, ATTN: PERI-I
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARI Research Note 92-29

11. SUPPLEMENTARY NOTES
Contracting Officer's Representative, Donald F. Haggard

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT (Maximum 200 words)

Tank gunnery training devices are designed to decrease the costs and other resources required for training. To realize the potential resource savings of devices and, at the same time, maintain desired proficiency, cost-effective tradeoffs must be specified for device training and on-tank training. This report presents methods for deriving and using these specifications. In addition, the report summarizes the support requirements for obtaining the data necessary for using the methods specified. The recommended approach is to obtain the data from actual performance-based gunnery data. At the same time, it was recognized that this may not be feasible. Because of this difficulty, an alternative judgment-based method, called "simulated transfer," is offered to generate surrogate performance data.

14. SUBJECT TERMS
Training devices
Training management
Military training
Gunnery trainers
Tanks (combat vehicles)

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

In a time of decreasing budgets and increasing training costs, the U.S. Army has adopted simulation as a cost-effective alternative to field training. Evaluating the relative effectiveness of different gunnery training methods (e.g., training devices, dry-fire gunnery, and live-fire gunnery) is a complex issue involving tradeoffs between resources available for training gunnery and their effects on gunnery proficiency. Research presented here is part of a project that will develop methods for performing research to determine those tradeoffs. A companion report (Hoffman and Morrison, 1991) provides a technical discussion of the methods for designing research and analyzing data. This report presents a user's introduction to the methods described in that report. These methods represent an important contribution to the exploratory development program of the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and will facilitate execution of well-directed tradeoff research for tank gunnery training.

This research, part of the ARI task entitled "Application of Technology to Meet Armor Skills Training Needs," was performed under the auspices of ARI's Armor Research and Development Activity at Fort Knox. The proponent for the research is the Deputy Chief of Staff, Training, of the Training and Doctrine Command. The requirement for this research has also been recognized by the Secretary of Defense.

EDGAR M. JOHNSON
Technical Director

A USER'S INTRODUCTION TO DETERMINING COST-EFFECTIVE TRADEOFFS AMONG TANK GUNNERY TRAINING METHODS

EXECUTIVE SUMMARY

Requirement:

Tank gunnery training devices are designed to decrease costs and other resources required for training. On the other hand, they are suspected of being deficient in their training capabilities when compared to training on the tank. To realize the potential resource savings of devices and, at the same time, maintain desired proficiency, cost-effective tradeoffs must be specified for device training and on-tank training. Research methods are required to generate those tradeoff specifications. This report introduces research methods that address this tradeoff problem.

Procedure:

The report begins with a discussion of the basic relationships underlying tradeoffs. Techniques for determining tradeoffs between two methods (the simplest case) are presented. These methods are expanded to more complex cases (three or more training methods). Some of the problems in obtaining performance data required by these methods are pointed out. Because of these problems, methods for obtaining surrogate performance data are also discussed.

Findings:

Several options, with varying degrees of sophistication, are specified for determining tradeoffs from actual performance data. All of the methods require the measurement or estimate of proficiency at multiple points during training. Further, in contrast to training transfer research, training and multiple assessments of proficiency must be considered on the operational equipment as well as on the alternative training devices. From these data, curvilinear functions can be derived to determine the amount of device training that minimizes training costs. Data collection requirements and data

analysis methods are intricate, particularly if more than two training methods are being considered. Consequently, tradeoff research is costly, and obtaining results is not guaranteed. Because of the difficulties of conducting tradeoff research, a judgment-based "simulated transfer" method is suggested for making tradeoff specifications in lieu of empirical data.

Utilization of Findings:

Policymakers and research planners have typically underestimated the effort needed to address training tradeoff questions. This research found no easy answers to the problem of determining resource tradeoffs among alternative training methods. Data collection and data analysis requirements are complex. To address tradeoff questions empirically, significant research resources must be committed to the effort. Even then, the focus may need to be kept fairly narrow.

A USER'S INTRODUCTION TO DETERMINING COST-EFFECTIVE TRADEOFFS AMONG TANK GUNNERY TRAINING METHODS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

A USER'S INTRODUCTION TO DETERMINING COST-EFFECTIVE TRADEOFFS AMONG TANK GUNNERY TRAINING METHODS

Tank gunnery training devices are designed to decrease the costs and other resources required for training. On the other hand, they are suspected of being deficient in their training capabilities compared with live-fire training on the tank. Too much reliance on tank training results in greater costs and too much reliance on the training device may hurt proficiency. To realize the potential resource savings of devices and, at the same time, maintain desired proficiency, cost-effective tradeoffs must be specified between investments in device training and on-tank training. Research methods are required to generate tradeoff specifications that indicate how much the device training should be used.

Policy makers and research planners typically underestimate the effort needed to address tradeoff questions. The purpose of the present report is to provide an appreciation for the requirements of tradeoff research in an armor context. This overview of the methods is intended for a nontechnical audience, including armor policy makers and training managers who are interested in the problem. A more technical presentation of these issues is presented by Hoffman and Morrison (1991).

Basic Relationships

The basic relationship underlying a tradeoff is the effect that one method of training has on another method. These methods might be two different approaches to training on the tank (e.g., dry- vs. live-fire) or two different training devices; for illustration purposes, however, the present section discusses the relationship between a training device and live-fire training on the tank where the cost-effectiveness tradeoffs are most striking. The determination of this tradeoff requires more than simply estimating the amount of training transfer that occurs from training on a device to performance with live-fire on the tank. There are three important reasons for a more complicated model of analysis.

First, tradeoff research requires consideration of the training effectiveness of both methods. Typical transfer of training research focuses on the training effectiveness of one method, with effectiveness defined by performance in the "transfer" setting, i.e., on the tank. Transfer is positive if device training improves on tank performance. Tradeoff research goes a step beyond. It concerns a comparison of training results from the two methods. Thus, data must be collected that (a) shows how performance improves as a result of device training and (b) shows how performance improves as a result of live-fire training. Only then can questions be addressed about which method is best to use for training.

Second, comparisons of results from one level of training may not apply at other levels. This is because practice does not typically result in a linear (i.e., straight-line) increase in performance. Rather, performance tends to improve in increasingly smaller amounts as illustrated by the familiar learning curve. For example, Figure 1 presents hypothetical learning curves for training on the tank and training on some device. Both methods show the typical learning curve shape for which proficiency increases relatively fast early in training. Improvement in proficiency continues with

additional training, but the increases in proficiency become smaller and smaller. The result is that after relatively large amounts of training, the amount of improvement with each trial is so small that performance appears to level out at or near a maximum value, called an asymptote. In our example, the only technical difference between the two learning curves is that tank training has a higher asymptote, thus indicating that there are aspects of gunnery that are not trained on the device. As a result, equipment training is better than device training for all amounts of practice, but this difference increases with greater amounts. In other words, the size of the difference between tank training and device training, in our hypothetical example, depends on how much training is given.

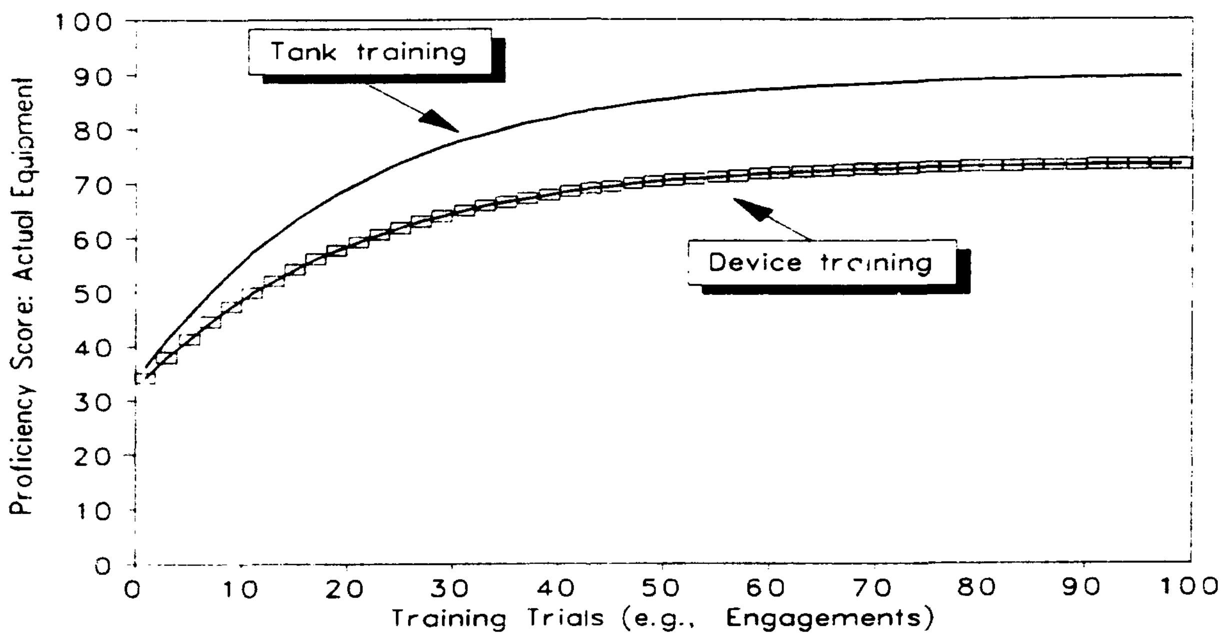

Figure 1. Hypothetical effects of training on the tank and on a gunnery training device as a function of practice trials.

Third, tradeoff research is concerned with more than just differences in performance. It also implies an assessment of the relative time or costs by which the performance differences are achieved. Thus, training with the tank may result in higher performance than training with the device alone; however, if training with the device is less expensive or less time consuming than training on the tank, there may be some economy in using the device for at least part of the training. Figure 2 illustrates what might happen if the hypothetical learning curves from Figure 1 are plotted by training time or cost rather than training trials. Figure 2 assumes that device training is one-fourth as time consuming or expensive as equipment training. Because the device gives more practice per unit of cost or time, there now appears to be an advantage to using the device early in training. However, as proficiency approaches the asymptote, device training may become a liability, particularly if there is a large difference between device and on-tank training in their respective asymptotes. Given this background, the tradeoff question becomes "How much device training is optimum to realize cost benefits but without sacrificing proficiency?" The next section examines research procedures for addressing this question.

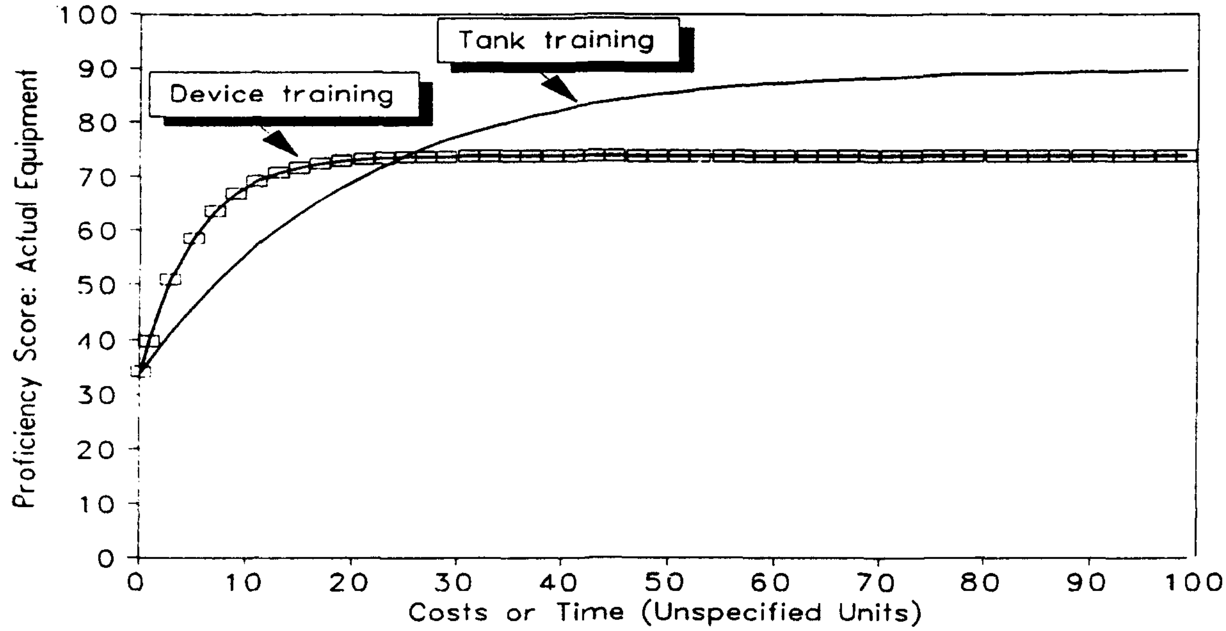

Figure 2. Hypothetical effects of training on the tank and on a gunnery training device as a function of training time or costs.

Determining Tradeoffs Between Two Methods

The relationships described above are based on hypothetical data. Ideally, tradeoff specifications should be based on actual performance data collected as tank crews train on both the device and on the tank. Plans for collecting performance data are often summarized as research designs. The following section presents examples of experimental research designs, so named because the research has explicit experimental control over the amount of training that crews receive. Hoffman and Morrison (1991) compared and contrasted experimental and nonexperimental designs for determining training tradeoffs, concluding that experimental designs are the more desirable alternatives. The principle advantage to experimental designs is that they allow the researcher to determine cause-effect relationships between training resources and proficiency.

The following sections describe three experimental designs that have been used as evidence that a potential tradeoff exists between training methods. As discussed below, the first two designs *do not* provide the information needed to determine the tradeoff specification. However, the two designs serve as heuristic introductions to the recommended third design.

Two-group, transfer-of-training design. In the following tables, the rows summarize the design requirements for each experimental condition (i.e., group of tank crews) as a sequence of training and testing events. As illustrated in Table 1, the classic transfer-of-training design stipulates that two groups be evaluated. The experimental group receives pretraining on the device, whereas the control group does not; both groups are then tested on the tank. The logic of the design is simple: Any performance differences between the two groups on the tank can be attributed to training on the device.

Table 1

Two-Group, Transfer-of-Training Design

	Event	
Group	Train	Test
Experimental	Device	Tank
Control	---	Tank

This simple research design addresses whether or not skills learned on the device transfer to performance on the tank, thereby indicating that the potential for a tradeoff exists. Recall from Figure 1, however, that proficiency varies nonlinearly as a function of the amount of device training. This design only generates two points on the device training curve--not enough to determine the shape of that function. In addition, it provides no information about how performance improves as a result of live-fire training.

<u>Multi-group, transfer-of-training design</u>. To generate more points on the device learning curve, additional experimental groups need to be added to the classic transfer-of-training design. As shown in Table 2, the experimental groups are differentiated by the amount of training that they

Table 2

Multi-Group, Transfer-of-Training Design

	Event			
Group	Train	Train	Train	Test
Experimental (3)	Device	Device	Device	Tank
Experimental (2)	---	Device	Device	Tank
Experimental (1)	---	---	Device	Tank
Control	---	---	---	Tank

receive on the device. This design assumes that a minimum of four data points are required to fit the relation describing tank performance as a function of device training: three points for estimating the curvilinear portion of the function and a fourth for estimating the asymptote. Nonlinear analysis

techniques can then be used to estimate the parameters of the function.[1] If proficiency on the device is assessed as training proceeds, the results from the groups that receive repeated device practice (Experimental 2 and 3) can also be used to determine the relationship of device proficiency as a function of practice.

A problem with this design, as well as the previous one, is that the effects of pretraining are evaluated only on initial performance on the tank. Such so-called "first shot" measurements of transfer may not show the ultimate training value of the device. In some cases, first-shot transfer may actually show negative transfer--that is, the experimental group(s) perform worse than the control group. If training on the criterion task continues, however, performance of the experimental group(s) may improve and actually overtake the control group, thereby showing that the device has a positive effect on training. The initial decrement in transfer performance is due to the effects of "interfacing" (Spears, 1985). Interfacing is caused by superficial differences between the device and the tank to which the crews must nevertheless adapt. Fortunately, the negative effects of interfacing on criterion performance are short-lived and are usually quickly overcome by the positive effects of device training. Because this design only assesses initial performance, however, the effects of device training could be significantly underestimated. Furthermore, as with the previous design, this one provides no information about how performance improves with live-fire training.

Groups-by-trials design. The problems with the previous designs can be corrected by requiring repeated practice on the tank so that most crews overcome any negative effects of interfacing and reach an acceptable standard of performance. In addition, this also allows an analysis of how performance improves with live-fire training. As illustrated in Table 3, this design is similar to the previous one in that it differentiates among groups who receive differing amounts of device training. The crucial difference is that all groups receive repeated training on the tank.[2] Thus, assuming that performance is continuously measured during the process of training on the tank, the immediate and the delayed effects of device training can be determined and the effects of live-fire training can be determined.

This design, then, provides the required data for a tradeoff analysis. Analysis of the data is begun by examining the tank performance data and setting a standard that most, if not all, crews have met by the end of training. Then, the performance of each crew is rescored as the number of training trials, time, or costs on the tank that were needed to reach the

---

[1]As discussed by Hoffman and Morrison (1991), there are complex and technical issues that pertain to using nonlinear estimation for fitting learning and transfer functions. To use these methods correctly for performing tradeoff research, one must be mindful throughout the research process of the demands of these methods, particularly of the demand for large samples.

[2]The four blocks of training illustrated in Table 3 may or may not provide enough repeated practice on the tank. The actual number of on-tank trials should be carefully determined in advance of data collection, either by pilot testing or in consultation with subject matter experts.

Table 3

Groups-by-Trials Design

Group	Event						
	Train	Train	Train	Train/ Test	Train/ Test	Train/ Test	Train/ Test
Experimental (3)	Device	Device	Device	Tank	Tank	Tank	Tank
Experimental (2)	---	Device	Device	Tank	Tank	Tank	Tank
Experimental (1)	---	---	Device	Tank	Tank	Tank	Tank
Control	---	---	---	Tank	Tank	Tank	Tank

standard. Using curve-fitting techniques, this variable is plotted as a function of training trials, time, or costs expended on the device as illustrated in Figure 3. This relation is sometimes called an iso-performance function, so-named because each point on the curve specifies a mix of training on the device and on the equipment that would result in a single standard of performance. In other words, this function describes the tradeoff of device and tank training for that performance *standard.*

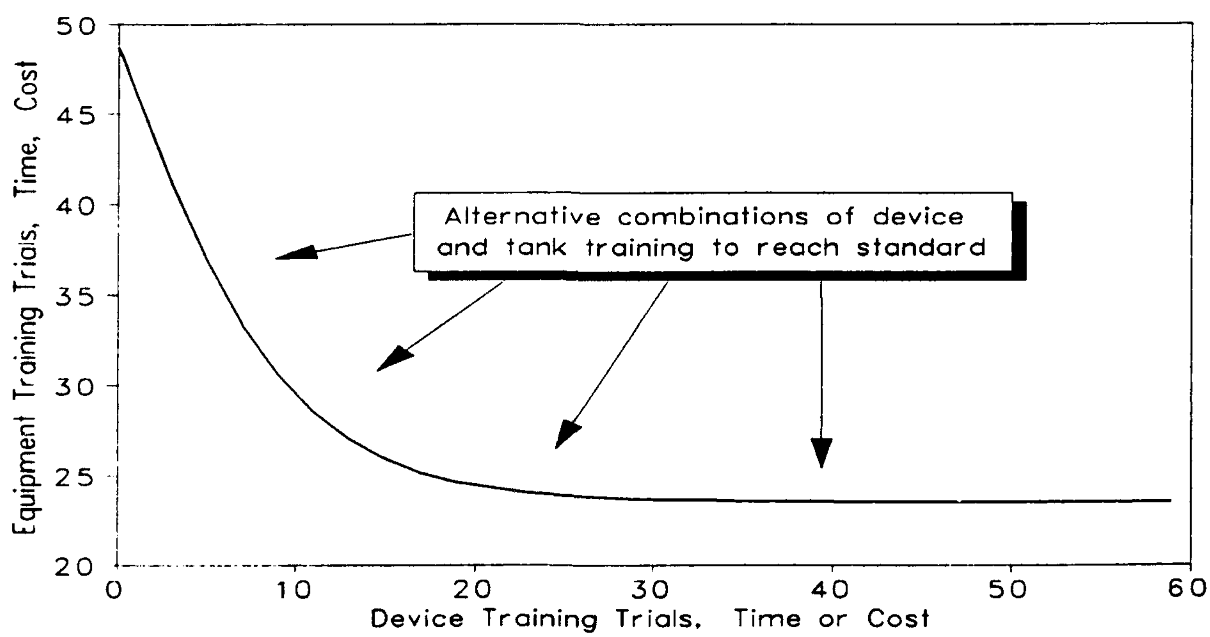

Figure 3. Iso-performance function generated from the hypothetical data.

The data used to plot the iso-performance function can also be rearranged to provide a total time or cost function. To derive this function, the total training time or costs to reach the performance standard must be

calculated for each crew. Note that these totals include training on both the device and the tank. These data are then plotted as a function of training time or costs that were expended on the device. As illustrated in Figure 4, the dip in the function is caused by the fact that device training is effective in reducing total training time or costs. The lowest point in the function identifies two important values. As projected on the vertical axis, this point corresponds to the minimum total time or costs that can be invested to reach the performance standard (about 39 time or cost units for the hypothetical data). On the horizontal axis, the dip identifies the amount of device training time or cost that results in the optimum investment of training resources (about 10 units for these data).

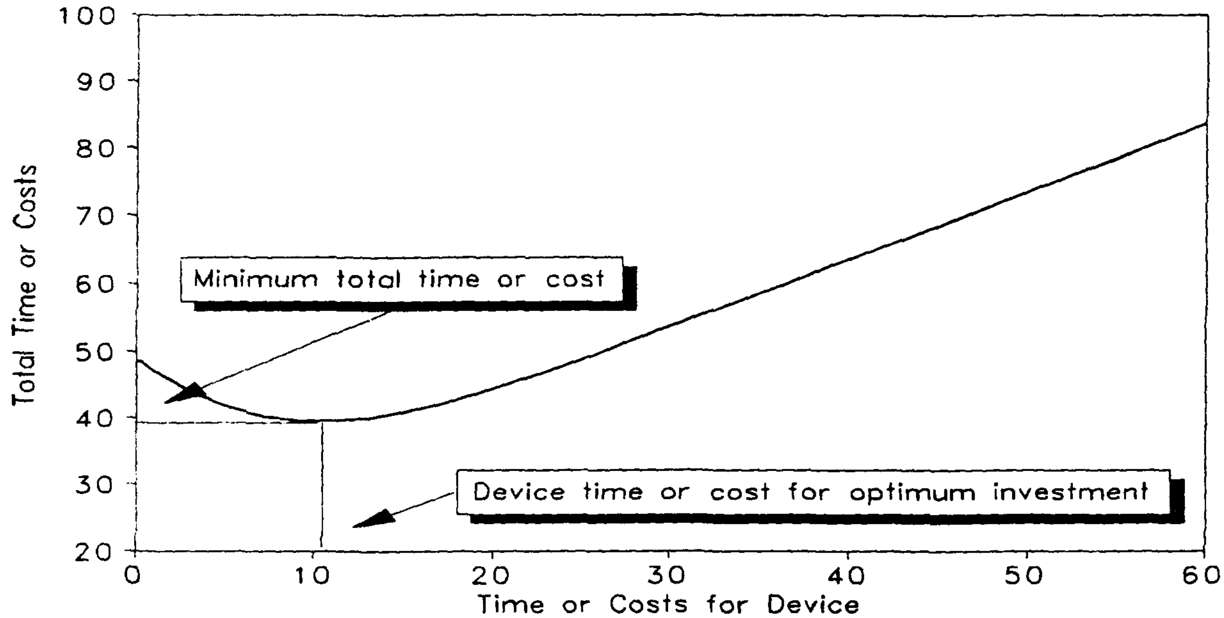

Figure 4. Total cost curve generated from the hypothetical data.

It is important to note that the optimum amount of device training is independent of the performance standard (Cronholm, 1985; Hoffman & Morrison, 1991). That is, the same amount of device training will be returned as the optimal value for any other performance standard. This fact implies that the results from this analysis are not specific to the performance standard set by the researcher. In fact, setting a performance standard in this sort of research is simply an intermediate step in data analysis and does not have to conform to established doctrine.

Determining Tradeoffs Among Three or More Methods

To this point, we have only considered the tradeoff between two methods: device training and live-fire training on the tank. However, tradeoffs involving more than two gunnery training methods are not difficult to imagine. For instance, a gunnery training program might specify that training begin on a low-fidelity, part-task trainer; proceed to a higher fidelity (and higher cost), whole-task trainer; and finish with live-fire training on the tank itself. The training developer's problem would then be to allocate time optimally among the three training methods. The techniques for determining

tradeoffs among three or more training methods are simply extensions of the two-method case. As discussed below, however, these multi-method problems introduce some serious practical complications that make the techniques difficult, if not infeasible, to implement in many cases.

Multi-dimensional designs. As shown in Table 4, the groups-by-trials design can be extended to the three-method example by factorially combining the four levels of training on both part- and whole-task training devices. This results in a design requiring 16 different groups, each of which receives repeated trials on the tank. This design allows the determination of all interactions among devices including the negative effects of interfacing, described earlier, as well as the positive synergistic effects of training on different devices (Morrison & Holding, 1990). Given sufficient data, this full factorial design can be used to construct three-dimensional, iso-performance and total cost curves analogous to the two-method case described above. For instance, the total cost function for this example would be a surface rather than a two-dimensional curve, with the lowest point in the surface describing the optimal allocation of training time to the part- and whole-task trainers.

As evident from Table 4, the most serious problem with the full factorial design is the sheer number of required experimental conditions. For all practical purposes, this design cannot be supported from a personnel or a logistical standpoint. Clearly, introducing more than three methods compounds this problem geometrically. One solution is to systematically eliminate some of those conditions from the design. For instance, fractional factorial designs can be constructed to reduce the number of required conditions by systematically confounding one or more of the effects that can be assessed in the full factorial design. For instance, an 8-group half fraction design can be constructed that confounds the interaction of the two devices (part- and whole-task trainers) but leaves the two device-by-tank training interactions interpretable as well as the three-way interaction. For higher level designs (i.e., using more than three methods), more exotic designs, such as latin square or response surface methods (e.g., Clark and Williges, 1973) can be used to reduce the number of required conditions. These reduced designs should be carefully constructed to conform with expectations about the interactions among training methods. Technical analysis support will be required at both the planning and data analysis phases of research to ensure that the results are interpretable.

Compromise method. Sticha, Schlager, Buede, Epstein, and Blacksten (1990) proposed an analytic solution for handling tradeoff problems involving more than two training alternatives which provides a compromise to the full factorial design. They suggested comparing cost function slopes instead of constructing iso-performance and total cost curves in order to identify optimum training mixes. Their expectation was that the point at which their learning curve slopes are equal will identify the same amount of device training identified by the minimum point on the total cost function (P. J. Sticha, personal communication, March 1991). That is, learning curves for several training methods could be adjusted to a common cost metric and the slopes of these curves compared to determine which method provided the greatest gain at various levels of proficiency.

Figure 5 illustrates the concept of comparing learning curve slopes to determine optimum training method. Three cost-based learning curves are

Table 4

Full Factorial Design for Determining Tradeoffs Among Three Methods

Group	Event									
	Train	Train	Train	Train	Train	Train	Train/ Test	Train/ Test	Train/ Test	Train/ Test
Experimental (3,3)	Part	Part	Part	Whole	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (2,3)	---	Part	Part	Whole	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (1,3)	---	---	Part	Whole	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (0,3)	---	---	---	Whole	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (3,2)	Part	Part	Part	---	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (2,2)	---	Part	Part	---	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (1,2)	---	---	Part	---	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (0,2)	---	---	---	---	Whole	Whole	Tank	Tank	Tank	Tank
Experimental (3,1)	Part	Part	Part	---	---	Whole	Tank	Tank	Tank	Tank
Experimental (2,1)	---	Part	Part	---	---	Whole	Tank	Tank	Tank	Tank
Experimental (1,1)	---	---	Part	---	---	Whole	Tank	Tank	Tank	Tank
Experimental (0,1)	---	---	---	---	---	Whole	Tank	Tank	Tank	Tank
Experimental (3,0)	Part	Part	Part	---	---	---	Tank	Tank	Tank	Tank
Experimental (2,0)	---	Part	Part	---	---	---	Tank	Tank	Tank	Tank
Experimental (1,0)	---	---	Part	---	---	---	Tank	Tank	Tank	Tank
Control (0,0)	---	---	---	---	---	---	Tank	Tank	Tank	Tank

Note. "Part" indicates a low-fidelity, part-task training device, whereas "whole" denotes a high-fidelity, whole-task training device. Numbers in parentheses indicate the number of blocks of practice the groups receive of the part-task and the whole-task trainers, respectively.

presented for Methods 1, 2, and 3, which respectively correspond to the part-task trainer, the whole-task trainer, and the actual tank in our ongoing example. The curve for Method 1 has the steepest initial learning gradient, and therefore provides the most cost efficient training up to the proficiency level indicated by point A. At point A, the slope of the Method 2 curve equals the slope of the Method 1 curve; beyond point A the slope to Method 2 is greater than Method 1. Between points A and B, the slope of the learning curve for Method 2 is also greater than the slope for Method 3. Therefore, between proficiency levels A and B, Method 2 provides the most cost efficient training. Beyond point B, the learning rate for Method 3 is greater that either Method 1 or 2. Therefore, if training were to begin for novice students, Method 1 should be used until the students reach point A. Assuming that Method 3 is used as the cost metric, the amount of training on Method 1 needed to reach point A is equal in dollar amount to approximately 2 units of Method 3 training. At this point training would shift to Method B, as

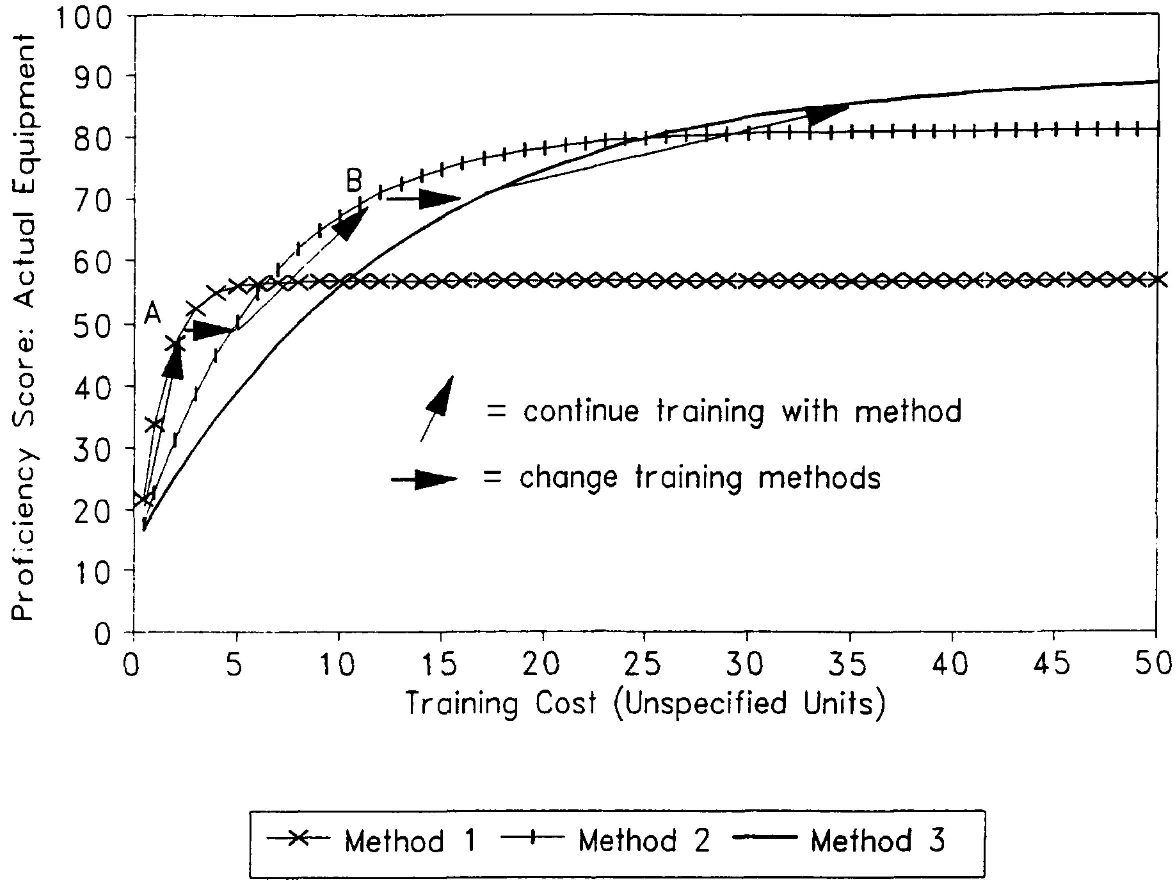

Figure 5. Training methods selected to maximize performance gain: Train with Method 1 to Point A, then with Method 2 to Point B, and finally with Method 3.

signified by the horizontal arrow, and students would now be progressing up the Method 2 learning curve. To improve from proficiency level A to level B would require Method 2 training that is equivalent to approximately 7 units of Method 3 training (12 - 5 on the X-axis). At level B, training shifts to Method 3, and students would be progressing up the Method 3 curve. Notice that the cost savings of combining methods are indicated in the figure as the total lengths of the two horizontal arrows. That is, by using the recommended sequencing Method 1 saves 3 units of cost in reaching 50% proficiency and Method B saves 8 units of cost in reaching 70% proficiency, for a total of 11 cost units saved over using Method 3 alone.

The figure also illustrates two other important points about training tradeoffs. The first point is that device optimization is independent of any performance standard. Regardless of the final proficiency target, costs are minimized by using Method 1 until proficiency reaches approximately 50%, followed by Method 2 up to a proficiency level of 70%, and then completing training on Method 3. The second point is related to the first: The allocation of time to training methods depends on the student as well as the characteristics of the learning curves. Using Figure 5 as an example, novice students should train on Method 1 for about 2 cost units worth of time or trials. On the other hand, more advanced students, who are more than 50% proficient, should not use Method 1 at all. Thus, questions about alternative

Table 5

Design for Obtaining Data Required for the Compromise Solution

Group	Event									
	Train	Train	Train	Train	Train	Train	Train/ Test	Train/ Test	Train/ Test	Train/ Test
Experimental (3,0)	Part	Part	Part	---	---	---	Tank	---	---	---
Experimental (2,0)	---	Part	Part	---	---	---	Tank	---	---	---
Experimental (1,0)	---	---	Part	---	---	---	Tank	---	---	---
Experimental (0,3)	---	---	---	Whole	Whole	Whole	Tank	---	---	---
Experimental (0,2)	---	---	---	---	Whole	Whole	Tank	---	---	---
Experimental (0,1)	---	---	---	---	---	Whole	Tank	---	---	---
Control (0,0)	---	---	---	---	---	---	Tank	Tank	Tank	Tank

Note. "Part" indicates a low-fidelity, part-task training device, whereas "whole" denotes a high-fidelity, whole-task training device. Numbers in parentheses indicate the number of blocks of practice the groups receive on the part-task and the whole-task trainers, respectively.

resource investments must always consider the background of population to be trained.

Constructing Figure 5 requires learning curves for each training method relating amount of device practice to proficiency on the tank. Table 5 illustrates the minimum design for obtaining this information. In essence, this table describes three experiments. Two experiments are first-shot, transfer-of-training experiments to determine the effects of the device training on tank performance. The third experiment, assesses the effects of tank training by specifying that crews in the control condition receive repeated practice on the actual equipment. Note that the performance of this group on the first practice trial on the tank also serves as the control (i.e., no device training) condition for the two first-shot transfer experiments. This design requires one less condition than the half fraction described above, but the most striking economy is that it requires repeated live-fire practice for only one group.

Although the design requires less support than the previous ones, it highlights the critical assumption of the compromise approach. The group-by-trials design was suggested because of the weakness of focusing on first-trial performance on the operational equipment following device training. The interpretation of the learning functions, depicted in Figure 5, assumes that learning on the operational equipment may be depicted by the same learning curve whether or not there is prior training on a device. As illustrated above, learning is depicted as movement along a learning curve. Switching training methods during the course of learning is modeled as switching from one learning curve to the other, *beginning on the operational equipment curve at the same performance level at which device pretraining was terminated.* These is no provision for "interfacing," which would be represented as performance on the operational equipment beginning below that expected from final device performance. Neither is there any provision for potential

synergistic interaction effects where training improvement based on combinations of methods are greater than any one method alone.

Notice also that the previous approach of directly estimating iso-performance curves from full factorial data avoids the problem. Any interactions among devices are incorporated in the estimates of the number of equipment trials that are needed to reach standard after device training has been completed. In terms of assumptions, the better solution is obtained from full factorial experiment requiring 16 experimental groups. A risky compromise solution is to assume that the effects of interactions among devices are minimal or non-existent, in which case only 7 experimental groups are required. Thus, choice of the full factorial versus the compromise design represents a tradeoff of their support requirements and in the realism of their initial assumptions.

Mixed training orders. The research designs described above assume that the alternative training methods are presented in a fixed order. In terms of both scheduling and training technique, this is not a very good assumption. The sustainment training concept is built on a continuous cycle that includes both live-fire and device training, so that the annual or semiannual cycle of training will intermingle training on devices and on the tank. In addition, within the gunnery cycle, not all training methods will be available at the same time for all crews. Some sharing of resources and alterations of training sequence must occur because of logistics. In terms of training technique, alternating between devices and on tank may reduce the negative effect of interfacing and may even promote a positive synergistic effect. Unfortunately, the problem of systematically describing tradeoff functions is further complicated when training alternates among methods. As a result, simplifying assumptions are required to even begin to address tradeoffs when training order is mixed. Some of the assumptions and a more complete discussion of this problem are provided in Hoffman and Morrison (1991).

Problems in Collecting Performance Data

The previous sections have demonstrated that a variety of useful techniques exist for determining tradeoffs among alternative training methods. At the same time, there are some significant impediments to collecting performance data on which those methods are based. Three impediments are particularly important for tradeoff research on tank gunnery.

Intervening events. Often researchers are required to conduct their investigations concurrently with ongoing training. As a result, the training method of interest is only one of several other methods being used to prepare crews and platoon for their gunnery qualification. For illustration, initial plans were drafted recently for examining the contribution of PRIME (Precision Range Integrated Maneuver Exercise) to platoon performance on Tank Table XII, the tank platoon qualification exercise. Figure 6 was constructed to provide a conceptual map of the training events that surround PRIME and Table XII. The boxes describe events or conditions that control performance. The dark boxes refer to platoon variables. The lighter, layered boxed refer to crew variables. The arrows between the boxes depict the relationships between the various training events and other variables that need to be determined in order to describe the training effectiveness of PRIME. The solid arrows indicate expected positive effects of training and training transfer. The

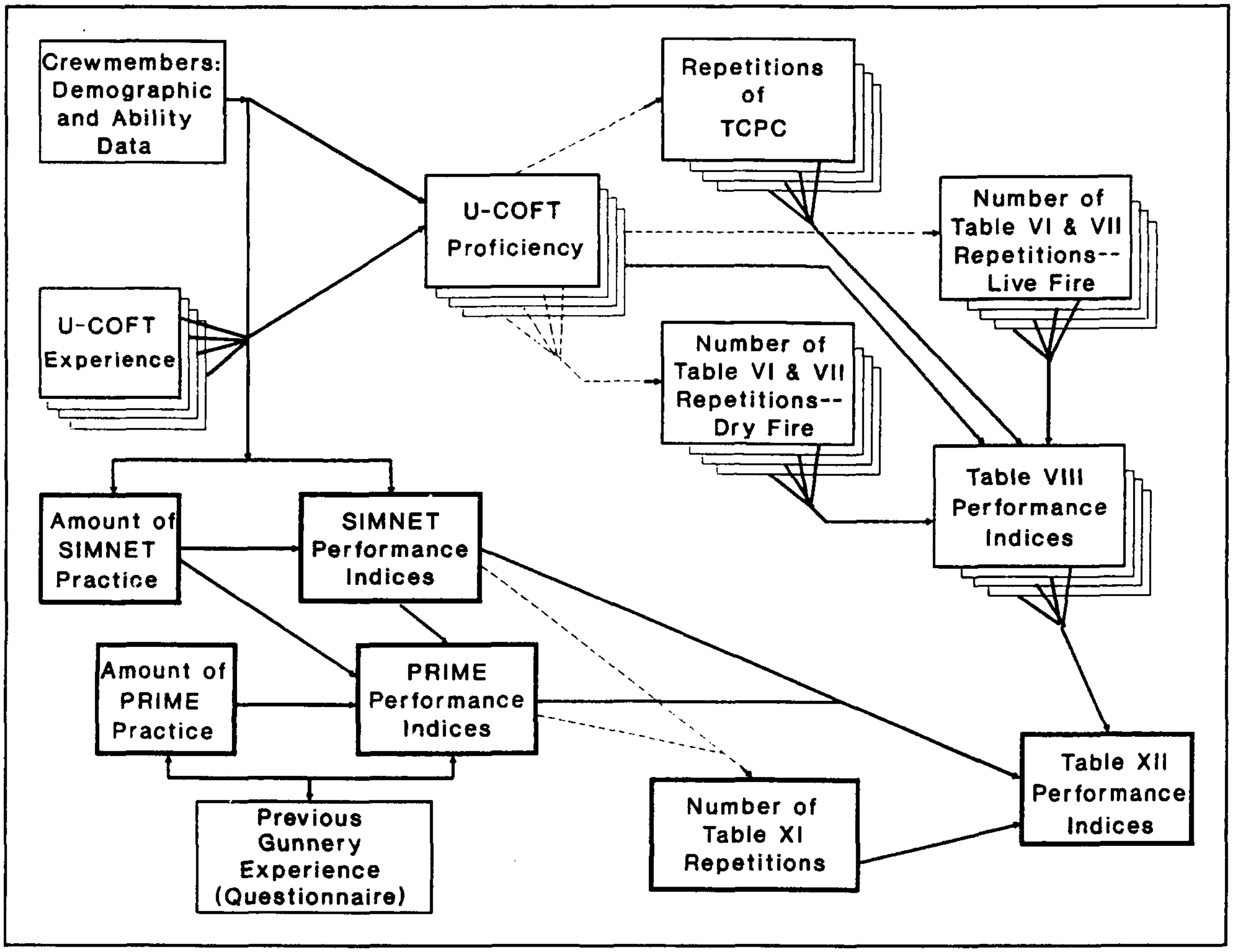

Figure 6. Sample map of training events and outcomes.

dotted arrows indicate expected negative effects created by training management decisions. For instance, less proficient crews are often allocated additional practice. The left-to-right flow of the figure presents the presumed chronological ordering of training events. The top portion of the figure centers on crew-level training and the bottom centers on platoon-level training.

From this figure, it is immediately obvious that there are numerous events which are expected to occur between PRIME training and Table XII. Even if the amount of PRIME were experimentally controlled, these events by themselves would act to cloud the relationship between PRIME training and Table XII performance. In addition, the management of some of these events would be expected to give training priority to the lower performing crews and platoons. This would tend to counteract any differences that manipulating PRIME training may create. As a result, any difference in gunnery proficiency produced by PRIME is likely to be diluted by the host of other training events.

<u>Performance criteria</u>. Generally, the crew- and platoon-level qualification exercises (Tables VIII and XII) are the performance criteria of

greatest interest for tank gunnery. Use of these doctrinal measures of gunnery performance poses several problems for tradeoff research.

First, tradeoff research assumes that live-fire gunnery is treated as just another method for training. This is quite different from traditional ways of thinking about gunnery training. Typically, each of the Combat Tables is treated as a separate training event that each crew must participate in once. If this approach is accepted as a given such that there is no freedom to consider changes to on-tank training, then there is no need for a tradeoff analysis. The addition of a device to the training program is just that--an addition. Without the possibility of reducing or increasing Tank Table training, there is no room for using a training device to create cost or performance optimization. Spending extra training dollars on device training may in fact increase performance. However, performance might be increased even more if a portion of those dollars were devoted to device training and the remainder were spent on additional on-tank training, or alternatively, if greater resources were spent on device training and less spent on training on the tank. It should be noted that such experimental changes to gunnery training would tend to meet considerable resistance, often thwarting efforts to conduct the kind of empirical research needed to answer the tradeoff question.

Second, live-fire performance, even if it could be manipulated for research purposes, tends to unreliable (e.g., Hoffman, 1989). That is, as a result of measurement difficulties and the inherent nature of the task, performance on any single engagement is at best only a rough indicator of proficiency. For example, luck appears to play a role in target acquisition, and round-to-round dispersion can cause misses even when sight pictures are correct. Increasing the number of engagements that are scored can increase reliability, but performance on a day or night portion of a gunnery table, each with five engagements, is still not very reliable. As result, documenting improvements in gunnery proficiency is difficult.

Third, because of safety restrictions and limitations on rounds, the tank tables do not test the full range of gunnery objectives (Hoffman, Fotouhi, Meade, & Blacksten, 1990). A training device may train more than can be tested on a live-fire range. For instance, the Unit Conduct-of-Fire Trainer (U-COFT) can train degraded modes of gunnery that cannot be fully assessed on a live-fire tank range. As result, the full potential of the training device is not evaluated by doctrinal performance measures.

Finally, live-fire training is time consuming and expensive. If permission were obtained, considerable resources would be required to increase the amount of live-fire training and experimentally manipulate how it is distributed among crews and platoons. The importance of the tradeoff question may not justify the time and expense needed to answer it with live-fire performance.

Sample size. Another roadblock to applying tradeoff research techniques is the number of crews or platoons that would be required for the research. The complexity of the research design, the intricacy of the data analysis, the number of intervening events, and live-fire criterion difficulties all combine to force consideration of sample sizes. We have estimated, that a brigade (two battalions, or approximately 100 tanks) is a bare minimum for crew-level research. Even then, relationships may be so clouded due to intervening

events and unreliable measurement, thereby reducing the probability of detecting real effects. Studying three or four battalions would be better from a statistical point of view. However, adding more units to the research adds to the number of potential intervening and extraneous events that could occur. Thus, larger sample sizes are better, but they are also likely to be accompanied by more disruptive influences, and they may exacerbate measurement problems.

Methods for Obtaining Surrogate Data

If at all possible, the researcher should seek to derive tradeoff relationships from empirical performance data. However, because of the problems and constraints discussed above, empirical performance-based research may be nearly infeasible. Research sponsors may not be willing to commit the time and resources required to answer seemingly simple questions about training tradeoffs. In this case, alternative judgment-based research methods may be appropriate for research-proficiency tradeoff questions. Hoffman and Morrison (1991) reviewed these methods and identified an appropriate method, which was referred to as "simulated transfer" by Pfeiffer and Horey (1988). Simulated transfer is a remarkably straightforward approach to predicting transfer: A subject matter expert (SME), who is familiar with performance both on the to-be-evaluated training device and on the operational equipment, is asked directly to estimate the data that would be obtained from a transfer experiment.

Hoffman and Morrison (1991) modified the simulated transfer procedure for application to tradeoff problems in gunnery training. They tested aspects of their modified approach by simply asking SMEs to respond to a structured questionnaire (Figure 7) intended to capture the learning curves for three gunnery training methods. As shown in the figure, performance estimates were provided for individual components of gunnery, such as searching for targets, fire commands, and tracking and switchology.

The method met with limited success at best: interrater agreement was high, and differences between methods and components conformed with conventional wisdom; however, the obtained learning functions were linear, and SMEs appeared to underestimate the number of practice trials that would be required to reach mastery (i.e., the asymptote). Hoffman and Morrison (1991) concluded that the simulated transfer approach may remain viable if further research is conducted to refine the procedures. For instance, research should be performed to compare and contrast results from alternative formats proposed for gathering simulated transfer data. Iterative feedback of results should be given to SMEs showing the implications of their ratings. Discussions among SMEs may indicate some of the implicit assumptions that each made about the process of transfer. Other research questions about the method include whether graphic feedback helps the raters to provide more reasonable (i.e., curvilinear) functions, and how quickly consensus can be reached on the shape and location of learning curves.

Note that extensive interventions in the data collection process begin to smack of data manipulation to reach conclusions that are preordained by the researcher. On the other hand, if empirical research cannot be performed, judgments are all that are available for making tradeoff decisions. These surrogate methods have the potential to structure and improve those judgments.

Training methods and tank gunnery components	What proficiency level (from 0 to 100) would you expect a novice crew to achieve after firing each of the following numbers of tasks in unit training? (Assume the tasks are similar to those in GUARD FIST I and the combat tables.)						What is the highest profi-ciency possible using only this training method?	How confident are you of your estimates? 1-Very 2-Quite 3-Somewhat 4-Slightly 5-Not at all
	0 Tasks	10 Tasks	20 Tasks	30 Tasks	40 Tasks	50 Tasks		
MILES (e.g., Tables II, III, and IV) ONLY								
Search								
Acquisition Reports								
Normal mode fire commands and reengagement								
Degraded mode fire commands and subsequent commands								
Spot reports								
GUARD FIST I ONLY								
Search								
Acquisition reports								
Reaction drills								
Normal mode fire commands and reengagement								
Degraded mode fire commands and subsequent commands								
Tracking and switchology								
Spot reports								
Live fire (e.g. Table VII or VIII) ONLY								
Search								
Acquisition reports								
Normal mode fire commands and reengagement								
Degraded mode fire commands and subsequent commands								
Tracking and switchology								

Figure 7. Questions from simulated research questionnaire.

REFERENCES

Clark, C., & Williges, R.C. (1973). Response surface methodology central-composite design modifications for human performance research. Human Factors, 15, 295-310.

Cronholm, J.N. (1985). The optimization of training systems. Proceedings of the Symposium on the Transfer of Training to Military Operational Systems (pp. 237-259). Brussels: North Atlantic Treaty Organization.

Hoffman, R.G. (1989). Description and prediction of Grafenwoehr M1 Table VIII performance (ARI Tech. Rep. 837). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B136 331)

Hoffman, R.G., Fotouhi, C.H., Meade, G.A., & Blacksten, H.R. (1990). Design of a threat-based gunnery performance test: Issues and procedures for crew and platoon tank gunnery (ARI Tech. Rep. 893). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 356)

Hoffman, R.G., & Morrison, J.E. (1991). Methods for determining resource and proficiency tradeoffs among alternative tank gunnery training methods (HumRRO Final Report in preparation). Alexandria, VA: Human Resources Research Organization.

Morrison, J.E., & Holding. (1990). Designing a gunnery training strategy (ARI Tech. Rep. 899). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 129)

Pfeiffer, M.G., & Horey, J.D. (1988). Analytic approaches to forecasting and evaluating training effectiveness (Tech. Rep. 88-027). Orlando, FL: Naval Training Systems Center.

Spears, W.D. (1985). Measurement of learning and transfer through curve fitting. Human Factors, 27, 251-266.

Sticha, P.J., Schlager, M., Buede, D.M., Epstein, K., & Blacksten, H.R. (1990). Designing skill-based training: Skill characterization and preliminary model requirements (ARI Tech. Rep. 897). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A227 237)